Ultrasound Obstet Gynecol 2017
Published online in Wiley Online Library (wileyonlinelibrary.com).

Letter to the Editor

External validation of ADNEX model for diagnosing ovarian cancer: evaluating performance of differentiation between tumor subgroups

I would like to thank Araujo[1] and Meys[2] and their colleagues for their validation studies of the ADNEX model. For women with at least one persistent adnexal mass who are scheduled for surgery, the ADNEX model can predict their risk of having one of five different types of adnexal mass: benign, borderline, Stage I ovarian cancer (OC), Stage II–IV OC and secondary metastatic OC[3]. The overall predicted risk of malignancy is obtained by adding the predicted risks of the latter four subgroups. Despite the fact that validation of a multinomial risk model is not straightforward[4], both research groups carried out robust and well-reported studies. I have, however, some observations regarding the methodology and results of their assessment of the performance of the ADNEX model in differentiating between tumor subtypes. To this end, I requested further data from the authors and I would like to express my gratitude for their co-operation.

The area under the receiver–operating characteristics curve (AUC) of the ADNEX model for discriminating between benign and all malignant tumors can be obtained by the standard approach, using the overall predicted risk of malignancy. Using this method, the reported AUC was 0.925 by Araujo *et al.*[1] and 0.93 by Meys *et al.*[2], which is in line with other validation studies of the ADNEX model[3,5].

Discrimination between tumor subgroups is of particular interest and several methods exist to calculate the AUC for models differentiating between two tumor subgroups, including the recommended 'conditional risk' method[4]. For example, to obtain the AUC for models differentiating between borderline *vs* Stage I OC, the conditional risk for Stage I OC is defined as the risk for Stage I OC divided by the sum of the risks for borderline and Stage I OC. Then, after excluding patients with histology other than borderline or Stage I OC, the AUC can be obtained by standard procedures using the conditional risk.

Araujo and colleagues computed subgroup AUCs using the overall risk of malignancy instead of the conditional risk. The limitation of this is that the overall risk of malignancy is the sum of the risks of all four malignant subgroups, therefore incorporating risks for subgroups that are not of interest. Furthermore, when one is, for example, interested in the AUC for borderline *vs* Stage I OC, the overall risk of malignancy incorporates the risk for both subgroups. In this case, it is not clear how the overall risk of malignancy can assess optimally discrimination between these two subgroups. Table 1 shows the AUCs for each paired subgroup comparison as reported by Araujo *et al.* and as obtained using the conditional risk method. AUCs for five of the 10 subgroup comparisons were higher using the conditional risk method and two were lower. On average, AUCs were higher by 0.02 with the conditional risk method.

Meys and colleagues used the risk for only one of the subgroups when assessing performance of the ADNEX model for differentiating between subgroups (e.g. risk for Stage I OC used when assessing discrimination between borderline tumors and Stage I OC). The limitation here is that two patients with the same risk for Stage I OC can have very different risks for a borderline tumor because there are a total of five subgroups. Table 2 shows the AUCs for each paired comparison as reported by Meys *et al.* and as obtained using the conditional risk method. AUCs were higher for nine of the 10 subgroup comparisons using the conditional risk method. The AUC for borderline *vs* Stage I OC showed a substantial increase from 0.60 to 0.79. On average, AUCs were higher by 0.06 with the conditional risk method.

Table 1 Areas under receiver–operating characteristics curves for differentiation between tumor subgroups as reported by Araujo *et al.*[1] and as recalculated after applying conditional risk method to same data

Tumor subgroups	*Araujo* et al.[1]	*Conditional risk method*	*Difference*
Benign *vs* borderline	0.83 (0.73–0.94)	0.80 (0.67–0.88)	−0.03
Benign *vs* Stage I OC	0.88 (0.78–0.98)	0.94 (0.86–0.97)	0.06
Benign *vs* Stage II–IV OC	0.99 (0.98–1.00)	0.99 (0.88–> 0.99)	0.00
Benign *vs* metastatic	0.97 (0.93–1.00)	0.97 (0.89–0.99)	0.00
Borderline *vs* Stage I OC	0.64 (0.44–0.84)	0.64 (0.43–0.81)	0.00
Borderline *vs* Stage II–IV OC	0.97 (0.93–1.00)	0.93 (0.79–0.98)	−0.04
Borderline *vs* metastatic	0.77 (0.60–0.95)	0.80 (0.57–0.92)	0.03
Stage I OC *vs* Stage II–IV OC	0.94 (0.87–1.00)	0.96 (0.83–0.99)	0.02
Stage I OC *vs* metastatic	0.64 (0.42–0.86)	0.71 (0.46–0.88)	0.07
Stage II–IV OC *vs* metastatic	0.89 (0.78–1.00)	0.95 (0.81–0.99)	0.06

Values in parentheses are 95% CI. OC, ovarian cancer.

Table 2 Areas under receiver–operating characteristics curves for differentiation between tumor subgroups as reported by Meys *et al.*[2] and as recalculated after applying conditional risk method to same data

Tumor subgroups	*Meys* et al.[2]	*Conditional risk method*	*Difference*
Benign *vs* borderline	0.81 (0.75–0.86)	0.88 (0.81–0.92)	0.07
Benign *vs* Stage I OC	0.87 (0.84–0.91)	0.92 (0.87–0.96)	0.05
Benign *vs* Stage II–IV OC	0.97 (0.94–0.99)	0.97 (0.94–0.99)	0.00
Benign *vs* metastatic	0.93 (0.89–0.96)	0.96 (0.89–0.99)	0.03
Borderline *vs* Stage I OC	0.60 (0.44–0.74)	0.79 (0.62–0.90)	0.19
Borderline *vs* Stage II–IV OC	0.87 (0.78–0.93)	0.92 (0.83–0.96)	0.05
Borderline *vs* metastatic	0.90 (0.77–0.97)	0.96 (0.86–0.99)	0.06
Stage I OC *vs* Stage II–IV OC	0.82 (0.71–0.90)	0.85 (0.74–0.92)	0.03
Stage I OC *vs* metastatic	0.72 (0.53–0.86)	0.76 (0.55–0.89)	0.04
Stage II–IV OC *vs* metastatic	0.67 (0.55–0.78)	0.71 (0.53–0.84)	0.04

Values in parentheses are 95% CI. OC, ovarian cancer.

In conclusion, discrimination between tumor subgroups using the ADNEX model was better than reported. In fact, results are similar to the validation results in the prospective ADNEX study[3]. Of course, given the low prevalence of some tumor types (borderline, Stage I OC, secondary metastasis) and heterogeneity between centers, variation in results between individual centers is to be expected.

B. Van Calster
Department of Development and Regeneration, KU Leuven, Leuven, Belgium
(e-mail: ben.vancalster@kuleuven.be)
DOI: 10.1002/uog.17391

References

1. Araujo KG, Jales RM, Pereira PN, Yoshida A, de Angelo Andrade L, Sarian LO, Derchain S. Performance of the IOTA ADNEX model in preoperative discrimination of adnexal masses in a gynecological oncology center. *Ultrasound Obstet Gynecol* 2017; **49**: 778–783.
2. Meys EMJ, Jeelof LS, Achten NMJ, Slangen BFM, Lambrechts S, Kruitwagen RFPM, Van Gorp T. Estimating risk of malignancy in adnexal masses: external validation of the ADNEX model and comparison with other frequently used ultrasound methods. *Ultrasound Obstet Gynecol* 2017; **49**: 784–792.
3. Van Calster B, Van Hoorde K, Valentin L, Testa AC, Fischerova D, Van Holsbeke C, Savelli L, Franchi D, Epstein E, Kaijser J, Van Belle V, Czekierdowski A, Guerriero S, Fruscio R, Lanzani C, Scala F, Bourne T, Timmerman D. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study. *BMJ* 2014; **349**: g5920.
4. Van Calster B, Vergouwe Y, Van Belle V, Looman CWN, Timmerman D, Steyerberg EW. Assessing the discriminative ability of risk models for more than two outcome categories: a perspective. *Eur J Epidemiol* 2012; **27**: 761–770.
5. Sayasneh A, Ferrara L, De Cock B, Saso S, Al-Memar M, Johnson S, Kaijser J, Carvalho J, Husicka R, Smith A, Stalder C, Blanca M, Ettore G, Van Calster B, Timmerman D, Bourne T. Evaluating the risk of ovarian cancer before surgery using the ADNEX model: a prospective multicentre external validation study. *Br J Cancer* 2016; **115**: 542–548.

